REMARKS

Reconsideration of the above referenced application in view of the following remarks is requested. Existing claims 1-18 remain in the application.

SUMMARY OF EXAMINER INTERVIEW

Upon the request from the Applicant, the Examiner and the Applicant held a teleconference at 2:00pm (EST) on September 20, 2005. The Applicant pointed out that the prior art references, alone or in combination, do not teach or suggest elements as claimed. The Examiner did not give a response to the Applicant's arguments during that meeting and asked for more time to review the prior art references. On September 22, 2005, the Applicant called the Examiner again at 2:00pm (EST). During that phone conversation, the Examiner asked the Applicant to file an after-final amendment presenting the Applicant's arguments and the Examiner will withdraw the final Office Action. The following are the Applicants' arguments.

ARGUMENT

Claim Rejections – 35 U.S.C. § 102

Claims 5-6, 10-12, and 16-18 are rejected under 35 U.S.C. 102(b) as being anticipated by Gorin et al. (US Patent No. 6,044,337) (hereinafter Gorin).

The Examiner asserted in the Office Action dated 06/09/05 that Gorin discloses a mechanism, method, and computer-readable medium encoded with a program for keyword based speech recognition, comprising a keyword spotting mechanism for detecting, using at least one acoustic model, at least one keyword from input speech

Application No.: 09/891,610	Page 9 of 16	Examiner: Vo, H.
Filed: 06/27/2001		Art Unit: 2655

data based on a keyword list; and a keyword based recognition mechanism for recognizing a word sequence using the at least one keyword, detected by the keyword spotting mechanism, based on a language model. The Examiner relied on col. 5, line 1 to col. 6, line 34 and Figure 4 of Gorin to make the above assertion. This assertion is erroneous. The cited text discloses a method for automatically generating and selecting superwords as well as meaningful phrases based on minimization of the language perplexity on a training corpus (see col. 3, line 11 to col. 5, line 20, and Figure 1); comparison data to show that the addition of superwords and meaningful phrases to the lexicon of a speech recognizer reduces false rejections (see col. 5, lines 21 to 29, and Figure 2)); a speech recognizer incorporating a meaningful phrase selector and a superword selector, which use the method to select superwords and meaningful phrases based on the training corpus (see col. 5, lines 30 to 42, and Figure 3); and a structure of superword and meaningful phrase generation subsystem (see col. 5, line 43 to col. 6, line 34, and Figure 4).

The cited text and figures as well as the entire Gorin reference do not teach or suggest the keyword spotting based speech recognition technology which is disclosed and claimed (claims 5, 10, and 16) in the above-identified application. The keyword spotting based speech recognition technology <u>detects</u> keywords in a predetermined keyword list, using acoustic models, from an input utterance and then relies mainly on the detected keywords to figure out the utterance based on language models (see paragraphs 32 and 36 of the specification). The keyword list includes words that are substantially significant for an application (e.g., "television" in an application for automated control of home appliances) (see paragraph 33 of the specification).

Application No.: 09/891,610   **Page 10 of 16**   Examiner: Vo, H.
Filed: 06/27/2001   Art Unit: 2655

Keywords are not word combinations that are so often spoken (i.e., superwords or meaningful phrases in Gorin) such as "I would like to." No training process based on minimization of the language perplexity on a training corpus is required for generating keywords. In marked contrast, such training process are specifically required by generating superwords or meaningful phrases (see Figure 1, col. 3, line 11 to col. 5, line 20, and Figure 4, col. 5, line 43 to col. 6, line 34 of Gorin). Therefore, keywords are not superwords or meaningful phrases as disclosed in Gorin.

Additionally, the keyword list is not the lexicon as disclosed in Gorin because the keyword list only includes keywords while the lexicon includes all possible words (including added superwords and meaningful phrases) that a speech recognizer can recognize. The integrated speech recognition mechanism as disclosed in the above-identified application does have a lexicon as any other speech recognizer. This lexicon is used by the graph decoder based speech recognition mechanism only and includes keywords in the keyword list (see Figure 1 and paragraphs 22 and 26 of the specification).

Moreover, the keyword based recognition mechanism as recited in claim 5 may recognize words other than keywords in the keyword lists based on the detected keywords in an input utterance and the language model (see paragraphs 36 and 37 of the specification, and Figure 5 (constrained language model 240 and constrained language 510)). As an example, Figure 6 of the above-identified application shows that constrained language 510 may be derived from the constrained language model 240 based on the keyword lists. The constrained language 510 includes many words other than keywords in the keyword lists. When any keyword(s) is detected in an input

Application No.: 09/891,610   Page 11 of 16   ·Examiner: Vo, H.
Filed: 06/27/2001   Art Unit: 2655

PAGE 16/21 * RCVD AT 10/5/2005 4:12:44 PM [Eastern Daylight Time] * SVR:USPTO-EFXRF-6/31 * DNIS:2738300 * CSID:503 684 3245 * DURATION (mm-ss):07-26

utterance, the keyword(s) is/are used to search the constrained language to generate an output word sequence (e.g., 125 as shown in Figure 5). In fact, the keyword based speech recognition process does not use a lexicon, which is only used by the graph decoder based speech recognition process). In marked contrast, Gorin performs normal speech recognition using a lexicon with added superwords and meaningful phrases once these superwords and meaning phrases are generated. In other words, Gorin treats superwords and meaningful phrases in the same way as a single word in the lexicon (see col. 2, line 66 to col. 3, line 10 of Gorin) and thus, the recognition process disclosed in Gorin must use the lexicon with added superwords and meaningful phrases. Therefore, the claimed keyword based recognition process is not taught or suggested by Gorin at all.

In the final Office Action dated 06/09/05, the Examiner responded to the above arguments presented by the Applicant. In the Examiner's response, the Examiner states, "In order to recognize the meaning of the input, the speech recognizer must first recognize all the input words that are semantically significant before assigning a meaning to the input phrase. This is the same as or equivalent to spotting keywords." This statement is incorrect. The Examiner confused the keyword spotting technique with a speech recognition technique. In a keyword spotting system, a predetermined list of keywords is provided. When an input utterance comes, each keyword in the predetermined keyword list is compared to segments of the input utterance based on the acoustic models of the keyword. The keyword is spotted from the input utterance if one segment of the input utterance matches the acoustic models of the keyword. In a speech recognition system, however, an input utterance is divided into many small

Application No.: 09/891,610   Page 12 of 16   Examiner: Vo, H.
Filed: 06/27/2001   Art Unit: 2655

PAGE 17/21 * RCVD AT 10/5/2005 4:12:44 PM [Eastern Daylight Time] * SVR:USPTO-EFXRF-6/31 * DNIS:2738300 * CSID:503 684 3245 * DURATION (mm-ss):07-26

pieces and each piece is compared with all possible acoustic models in the system. Those acoustic models that produce the highest probability of matching the piece of the input utterance are selected as candidate models for this piece of the input utterance. The lexicon and language models (e.g., grammars) are further used to decide among candidate models among different pieces of the input utterance to produce one or more candidate sentences or phrases for the input utterance. Thus, spotting keywords and recognizing an utterance are different and spotting keywords is not the same or equivalent to recognizing an utterance, as asserted by the Examiner. Therefore, the Examiner failed to rebut the Applicant's arguments in the Examiner's response.

Because the Gorin reference does not teach or suggest any limitations recited in claims 5, 10, and 16, these claims are not anticipated by Gorin and are thus allowable over Gorin.

Because independent claims 5, 10, and 16 are allowable, all claims dependent therefrom are also allowable over Gorin (e.g., existing claims 6, 11-12, and 17-18).

Claim Rejections – 35 U.S.C. § 103

Claims 1-4, 7-9, and 13-15 are rejected under 35 U.S.C. 103(a) as being unpatentable over Hedin et al. (US Patent No. 6,185,535) (hereinafter (Hedin) in view of Setlur et al. (US Patent No. 5,956,675) (hereinafter Setlur), and further in view of Gorin et al. (US Patent No. 6,044,337) (hereinafter Gorin).

As to claim 1, the Examiner asserted that Hedin discloses a system comprising a first speech recognition mechanism for recognizing a word sequence from input speech data, based on a language model; and a second speech recognition mechanism for

Application No.: 09/891,610   Page 13 of 16   Examiner: Vo, H.
Filed: 06/27/2001   Art Unit: 2655

recognizing, when the first speech recognition mechanism fails, the word sequence based on at least one word detected from the input speech data. Specifically, the Examiner cited col. 5, lines 1 to 33 of Hedin to support his assertion. This assertion is erroneous. Although Hedin does disclose two speech recognition systems (a simple one located at a client side and a more powerful one located at a server side, see col. 4, line 66 to col. 5, line 3), Hedin does not teach or suggest the relationship of the two recognition systems recited in claim 1. Claim 1 in the above-identified application specifically recites that the second speech recognition mechanism is only activated when the first speech recognition mechanism fails. In other words, only one output result is generated by either the first speech recognition mechanism or the second recognition mechanism. In marked contrast, in Hedin, the simple automatic speech recognizer (ASR) at the client side recognizes some words (although in a very limited number) from an input utterance and are acted upon by the client (see col. 5, lines 5 to 17); and the more powerful ASR further recognizes those parts of the input utterance that are not recognized by the simple ASR (see col. 5, lines 18 to 23). In other words, the two ASRs in Hedin each generate a recognized result which is further acted upon by either the client or the server. Additionally, the entire Hedin reference does not disclose anything about the language model recited in claim 1. Therefore, Hedin does not teach or suggest all of the elements that the Examiner asserted it teaches, e.g., the relationship between the first and the second speech recognition mechanisms and the language model as recited in claim 1.

The Examiner asserted that Gorin teaches a keyword based speech recognition mechanism. For similar reasons presented for traversing 35 U.S.C. § 102 rejections as

Application No.: 09/891,610
Filed: 06/27/2001   Page 14 of 16   Examiner: Vo, H.
Art Unit: 2655

PAGE 19/21 * RCVD AT 10/5/2005 4:12:44 PM [Eastern Daylight Time] * SVR:USPTO-EFXRF-6/31 * DNIS:2738300 * CSID:503 684 3245 * DURATION (mm-ss):07-26

above, Gorin does not teach or suggest a keyword based speech recognition mechanism.

The Examiner asserted that col. 4, line 48 to col. 5, line 23 of Setlur teaches a graph-decoder based speech recognition mechanism. The cited text of Setlur does not teach graph-decoder based speech recognition. It only teaches a Viterbi-algorithm based speech recognition process, which is not the same as a graph-decoder based speech recognition process. Even if the Viterbi-algorithm based speech recognition process is the same as the graph-decoder based speech recognition process, a combination of Hedin, Setlur, and Gorin does not teach or suggest all the limitations recited in claim 1 because Hedin does not teach or suggest the relationship between the two speech recognition mechanisms and the language model recited in claim 1 and Gorin does not teach the keyword based speech recognition mechanism recited in claim 1. Therefore, claim 1 is allowable over Hedin in view of Setlur, and further in view of Gorin.

Regarding independent claims 7 and 13, the Examiner rejected them based on the same reasons as those used for rejecting independent claim 1. These claims recite similar limitations as those recited in claim 1. For similar reasons presented for traversing the Examiner's rejections to claim 1, the combination of Hedin, Setlur, and Gorin does not teach or suggest all the limitations recited in independent claims 7 and 13 and these two independent claims are thus allowable over Hedin in view of Setlur, and further in view of Gorin.

Application No.: 09/891,610   Page 15 of 16   Examiner: Vo, H.
Filed: 06/27/2001   Art Unit: 2655

PAGE 20/21 * RCVD AT 10/5/2005 4:12:44 PM [Eastern Daylight Time] * SVR:USPTO-EFXRF-6/31 * DNIS:2738300 * CSID:503 684 3245 * DURATION (mm-ss):07-26

Because independent claims 1, 7, and 13 are allowable over Hedin in view of Setlur, and further in view of Gorin, all claims dependent therefrom (e.g., existing claims 2-4, 8-9, and 14-15) are also allowable over the same combination of references.

CONCLUSION

In view of the foregoing, claims 1-18 are all in condition for allowance. If the Examiner has any questions, the Examiner is invited to contact the undersigned at (503) 264-1700. Early issuance of a Notice of Allowance is respectfully requested.
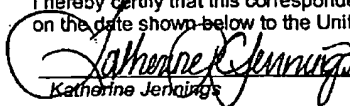
Respectfully submitted,

/Guojun Zhou/
Guojun Zhou
Patent Agent
Intel Corporation
Registration No. 56,478
(503) 264-1700

c/o Blakely, Sokoloff, Taylor & Zafman, LLP
12400 Wilshire Boulevard
Seventh Floor
Los Angeles, CA 90025-1026

Application No.: 09/891,610   Page 16 of 16   Examiner: Vo, H.
Filed: 06/27/2001   Art Unit: 2655